



1600

RAW SEQUENCE LISTING

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:41

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw

4 <110> APPLICANT: SANTI, Daniel
 5 PECK, Lawrence
 6 KEALEY, James
 7 DAYEM, Linda
 9 <120> TITLE OF INVENTION: Heterologous Production of Polyketides
 12 <130> FILE REFERENCE: 010025.01
 14 <140> CURRENT APPLICATION NUMBER: 09/699,136A
 15 <141> CURRENT FILING DATE: 2000-10-27
 17 <150> PRIOR APPLICATION NUMBER: 60/161,703
 18 <151> PRIOR FILING DATE: 1999-10-27
 20 <160> NUMBER OF SEQ ID NOS: 6
 22 <170> SOFTWARE: FastSEQ for Windows Version 4.0
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 1917
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Propionibacterium freudenreichii ssp.shermanii
 29 <400> SEQUENCE: 1

ENTERED

```

30 atgagcagca cggatcaggg gaccaacccc gccgacactg acgacctcac tcccaccaca 60
31 ctcagcctgg ccggggattt cccaaggcc actgaggagc agtgggagcg cgaagttgag 120
32 aaggtactca accgtggtcg tccaccggag aagcagttga ctttcgccga gtgtctgaag 180
33 cgcctgacgg ttcacaccgt cgatggcatc gacatcgtgc cgatgtaccg tccgaaggac 240
34 gccccgaaga agctgggtta ccccggcgtc gcacctttca cccgcgcac cacggtgcgc 300
35 aacggcgaca tggatgcctg ggacgtgcgc gccctgcacg aggatcccga cgagaagttc 360
36 acccgcaagg cgatcctcga aggcctggag cgtggcgtca cctccctggt gctgcgcgtt 420
37 gatcccgacg cgatcgaccc cgagcacctc gacgaggtcc tctccgacgt cctgctggaa 480
38 atgaccaagg tggaggtctt cagccgctac gaccaggggt ccgccgccga ggccctggtg 540
39 agcgtctacg agcgtctcga caagccggcg aaggacctgg ccctcaacct gggcctggat 600
40 cccatcgctg tcgcagccct gcagggcacc gagccggatc tgaccgtgct cggtgactgg 660
41 gtgcgccgcc tggcgaagtt ctgcgccgac tcgcgcgccg tcacgatcga cgcgaacatc 720
42 taccacaacg ccggtgcccg cgacgtggca gagctcgctt gggcactggc caccggcgcg 780
43 gagtacgtgc gcgccctggt cgagcagggc ttcaccgcca ccgaggcctt cgacacgatc 840
44 aacttccgtg tcaccgccac ccacgaccag ttcctcacga tcgcccgctt tcgcgccctg 900
45 cgcgaggcat gggcccgcac cggcgaggtc ttcggcgtgg acgaggacaa gcgcggcgcc 960
46 cgccagaatg cgatcaccag ctggcgtgac gtcacgcgcg aagaccctta tgtcaacatc 1020
47 cttcgcggtt cgattgccac cttctccgcc tccgttggtg gggccgagtc gatcacgacg 1080
48 ctgcccttca cccaggccct cggcctgccg gaggacgact tcccgtgctg catcgcgcg 1140
49 aacacgggca tcgtgctcgc cgaagaggtg aacatcggcc gcgtcaacga cccggccggt 1200
50 ggtcctact acgtcgagtc gtcacccgcg agcctggccg acgcccctg gaaggaattc 1260
51 caggaggtcg agaagctcgg tggcatgctg aaggccgtca tgaccgagca cgtcaccaag 1320
52 gtgctcgacg cctgcaatgc cgagcgcgcc aagcgccgtg ccaaccgcaa gcagccgatc 1380
53 accgcggtca gcgagttccc gatgatcggg gcccgagca tcgagaccaa gccgttcccc 1440
54 gccgctccgg cgcgcaaggg cctggcctgg catcgcgact ccgaggtgtt cgagcagctg 1500
55 atggatcgct ccaccagcgt ctccgagcgc cccaaggtgt tcctggcctg cttgggcacc 1560

```

RAW SEQUENCE LISTING

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:41

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw

```

56 cgtcgcgact tcggtggccg cgagggtctt tcgagcccg tgtggc'acat cgccggcatc 1620
57 gacaccccg aggtcgaagg cggcaccacc gccgagatcg tcgaggcatt caagaagtcg 1680
58 ggcgcccagg tggccgacct ctgctcgtcc gccaaagtct acgcgcagca gggacttgag 1740
59 gtcgccaagg cactcaaggc cgccggcgca aaggccctgt acctgtcggg cgccctcaag 1800
60 gagttcggtg atgacgccgc cgaggccgag aagctgatcg acggacgcct gtttatgggc 1860
61 atggatgtcg tcgacaccct gtcctccacc cttgatattc tgggagtcgc gaagtga 1917
63 <210> SEQ ID NO: 2
64 <211> LENGTH: 2187
65 <212> TYPE: DNA
66 <213> ORGANISM: Propionibacterium freudenreichii ssp. shermanii
68 <400> SEQUENCE: 2
69 gtgagcactc tgccccgttt tgattcagtt gacctcggca atgccccggt tcctgctgat 60
70 gccgcacgac gcttcgagga actggccgcc aaggccggca ccgagagagg gtgggagacg 120
71 gccgagcaga ttccggttgg caccctgttc aacgaagacg tctacaagga catggactgg 180
72 ctggacacct acgcaggtat cccgccgttc gtccacggcc cgtatgcaac catgtacgcg 240
73 ttccgtccct ggacgattcg ccagtacgcc ggtttctcca cggccaagga gtcgaacgcc 300
74 ttctaccgcc gcaaccttgc ggccggccag aaggccctgt cggttgccct cgacctgccc 360
75 acccaccgtg gctacgactc ggacaatccc cgctcgccg gtgacgtcgg catggccggt 420
76 gtggccatcg actccatcta tgacatgcgc gagctgttcg ccggcattcc gctggaccag 480
77 atgagcgtgt ccatgaccat gaacggcgcc gtgctgccga tcctggccct ctatgtggtg 540
78 accgccgagg agcagggcgt caagcccgag cagctcgccg ggacgatcca gaacgacatc 600
79 ctcaaggagt tcatggttcg taacacctac atctaccgcg cgacgccgag tatgcgaatc 660
80 atctctgaga tcttcgccta cacgagtgcc aatatgccga agtggaaattc gatttccatt 720
81 tccggtacc acatgcagga agccggcgcc acggccgaca tcgagatggc ctataacctg 780
82 gccgacggtg ttgactacat ccgcgccggc gactcgggtg gcctcaatgt cgaccagttc 840
83 gcgccgcgtc tgccttctt ctggggcatc ggcatgaact tcttcatgga ggttgccaag 900
84 ctgcgtgccg cgcgcattgt gtgggccaag ctggtgcac agttcgggcc gaagaacccg 960
85 aagtcgatga gcctgcgcac ccactcgag acctccggtt ggtcgtgac cgccaggagc 1020
86 gtctacaaca acgtcgtgcg tacctgcac gaggccatg ccgccacca gggccatacc 1080
87 cagtcgctgc acacgaactc gctcgcagag gccatcgccc tgccgaccga tttcagcgcc 1140
88 cgcacgccc gtaacaccca gctgttctc cagcaggaat cgggcacgac gcgcgtgatc 1200
89 gaccctgga gccgctcgcc atacgtcag gagctcacct gggacctggc ccgcaaggca 1260
90 tggggtcaca tccaggaggt cgagaaggtc ggccgcatg ccaaggccat cgaaaagggc 1320
91 atccccaaga tgcgcacgga ggaagccgcc gcccgacccc aggcacgcat cgactccggc 1380
92 cgccagccgc tgatcggcgt gaacaagtac cgctggagc acgagccgcc gctcgtatgt 1440
93 ctcaagggtg acaactccac ggtgctcgcc gagcagaagg ccaagctggt caagctgcgc 1500
94 gccgagcgcg atcccagaa ggtcaaggcc gccctcgaca agatcacctg ggccgcccgc 1560
95 aaccccgcag acaaggatcc ggatcgcaac ctgctgaagc tgtgcatcga cgctggccgc 1620
96 gccatggcga cggtcggcga gatgagcgac gcgctcgaga aggtcttcgg acgtacacc 1680
97 gccagattc gcaccatctc cgtgtgttac tcgaaggaag tgaagaacac gcctgagggt 1740
98 gaggaagcac gcgagctcgt tgaggaattc gagcaggccg agggccgtcg tcctcgcatc 1800
99 ctgctggcca agatgggcca ggacggtcac gaccgtggcc agaaggtcat cgccaccgcc 1860
100 tatgccgacc tcggtttcga cgtcgacgtg ggcccgctgt tccagacccc ggaggagacc 1920
101 gcacgtcagg ccgtcgaggc cgatgtgcac gtggtggcg tttcgtcgt cgccggcggg 1980
102 catctgacgc tggttccggc cctgcgcaag gagctggaca agctcggacg tcccacatc 2040
103 ctcatcaccg tggcgcgct gatccctgag caggacttcg acgagctgcg taaggacggc 2100
104 gccgtggaga tctacacccc cggcacccgtc attccggagt cggcgatctc gctggtcaag 2160
105 aaactcgagg ctctcgtcga tgcctag 2187
107 <210> SEQ ID NO: 3

```

RAW SEQUENCE LISTING

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:41

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw

```

108 <211> LENGTH: 1851
109 <212> TYPE: DNA
110 <213> ORGANISM: Streptomyces cinnamonensis
112 <400> SEQUENCE: 3
113 atgacggtcc tgcctgacga cgggctttcc ctggccgccc agttccctga tgcgacccat 60
114 gagcagtgcc accgccttgt ggaaggcgtc gtgcgcaagt cgggtaagga cgtatcgggg 120
115 acggccgcag aggaagcgct gtccaccacc ctcgaggacg ggctcaccac ccgccccctg 180
116 tacaccgcgc gcgacgccgc gcccgacgcc ggtttccccg gcttcgcccc ttctcgtcagg 240
117 ggttccgtcc ccgagggcaa caccgcgggc ggctgggacg tccggcagcg gtacgcgagc 300
118 gccgatcccc cgcgcaccaa tgaagcggtc ctcaccgatc tggagaacgg cgtcacctcg 360
119 ctctggctga cgctgggctc cgccggtctt ccggtcaccg gtctggagcg tgcgctcgac 420
120 ggctctacc tcgacctggt gcccgctcgc ctcgacgcgg gcagcgaggc cgcgaccgcc 480
121 gcccgggagt tgctgcgcct gtacgaggcc gcgggctcgc ccgacgacgc ggtgcgcggc 540
122 acgtcgcggc ccgaccgcgt cggccacgag gcccgaccg gggagaagag cacctccttc 600
123 gcggcggtcg ccgaactggc ccggtctgtc ggggagcgct accccggtct gcgcgcgctg 660
124 accgtggacg cgctgccgta ccacgaggcg ggtgcctccg ccgcgcagga gctcggcgcc 720
125 tcgctcgcca ccggtgtcga gtatctgcgc gccctgcacg acaagggtct cggtgtcgag 780
126 aaggccttcg cgcagctgga gttcagggtc gcggccaccg ccgaccagtt cctcaccatc 840
127 gccaaactgc gcgcgcgcgc ccgcctctgg gcgcgtgtcg ccgaggtgtc cggggtgccg 900
128 gccgcggggg cgcagcggca gcacgcggtg acctcgccgg tgatgatgac ccgccgcgac 960
129 ccgtgggtga acatgctgcg gaccaccgtg gcgtgcctcg gcgcgggtgt gggcggtgcc 1020
130 gacgccgtca cggctctgcc gttcgaccac gagctcgggc tgccggacgc gttcgcgcg 1080
131 cgcacgccc gcaacacgtc gacgatcctc ctcgaggagt cgcacctcgc gcgtgtcatc 1140
132 gaccggcgcg gcggctcctg gtacgtggag cgggtcaccg atgaactcgc ccacgcggcc 1200
133 tgggacttct tcaaggagat cgagcgcgcg gacggtcagg tcgcggcgct ccgctccggc 1260
134 ctggtcggcg accggatcgc cgcgacctgg gccgagcgca ggaagaagct ggcccggcgc 1320
135 cggaaccga tcacgggtgt cagcgagttc ccgctgtca ccgagcggcc ggtcgagcgc 1380
136 gagcccgcg ccgcgcggcc gcccgcggtt ctgcccgggg tgcgccgca cgaggcgta 1440
137 gaggagctgc gcggcgctc ggacgcgcac ctggaagcga ccggcgccc cccgaagggt 1500
138 ttcacgcgg cgctggggcc ggccgcgcgc 'cacacggcgc gcgcgacgtt cgccgccaac 1560
139 ctcttcacg cggcgcggtg cgagccgggt cacgaccgg tgtcgggtga cgcggagacg 1620
140 gccgcgagg cgttcgcggc gagcgcgcc accgtcgcgt gcctctgctc cagcgacgtg 1680
141 ctctacgcc agcaggcgga agcggtcgcg cgggccctga agtcggcggg cgcgctgcgg 1740
142 gtgttctcg cggggcgcg ggagttcgcc gacatcgacg agtacgtctt cgcgggctgc 1800
143 gacgcggtcg cgggtgtcac ctccaccctc gaccgcatgg gagtggcgta a 1851
145 <210> SEQ ID NO: 4
146 <211> LENGTH: 2202
147 <212> TYPE: DNA
148 <213> ORGANISM: Streptomyces cinnamonensis
150 <400> SEQUENCE: 4
151 atgcggatcc ccgaattcga cgacatcgaa ctgggcgcgg gcggcgggcc gtccggctcg 60
152 gcggagcagt ggcgcgcgc cgtgaaggag tccgtcggca agtccgagtc cgacctcctg 120
153 tgggagacgc ccgagggcat cgcggtcaag ccgctgtaca cgggcgcgca cgtcgagggc 180
154 ctggacttcc tggagacgta cccgggtgtc gcgcggtatc tgcgcgggcc ctacccgacg 240
155 atgtacgtga accagccgtg gacgatccgg cagtacgcgg gattctccac cgccgaggag 300
156 tccaacgcct tctaccgcgc caacctcgcg gcaggccaga aggggtcttc ggtcgccttc 360
157 gacctgcccc cgcaccgcgc gtacgacagc gaccaccgc gcgtcaccg tgacgtcggc 420
158 atggcgggcg tggccatcga ctccatctac gacatgcgtc agctcttcga cggcattccg 480
159 ctggacaaga tgacggtgtc gatgacgatg aacggtgccg tgctgcccg tctcgcgctg 540

```

RAW SEQUENCE LISTING

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:41

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw

```

160 tacatcgtgg cggccgagga gcagggcgta ccgcccggaga agctggcggg gaccattcag 600
161 aacgacatcc tcaaggagtt catggtccgc aacacctaca tctatccgcc gaagccctcg 660
162 atgcggatca tctccgacat cttcgcgtag acgtcgcgaga agatgccgcg ctacaactcc 720
163 atctcgatct ccggctatca catccaggag gcgggcgcgga cggccgacct ggagctggcg 780
164 tacacgctcg ccgacggtgt ggagtacctg cgagccgggc aggagggcggg cctggacgtg 840
165 gacgcgttcg cgccgcgggt ctccttcttc tgggcgatcg gcatgaactt cttcatggag 900
166 gtcgccaagc tccgcgcggc gcgcctgctc tgggcgaagc tcgtgaagca gttcgacctg 960
167 aagaacgcca agtccctctc cctgcgcacc cattcgcgaga catcgggctg gtcgctgacc 1020
168 gcgcaggacg tggtcaacaa cgtcacgcgc acgtgtgtcg aggcgatggc ggcgacgcag 1080
169 ggccacacgc agtccctgca cacgaacgcc ctggacgagg cgctcgccct gccgaccgac 1140
170 ttctccgcgc ggatcgcccg caacaccag ctgctcatcc agcaggagtc ggggacgacg 1200
171 cgacgatcg accgtgggg cggcagcgcg tacgtcgaga agctgacgta cgacctggcg 1260
172 cggcgcgctt gcgacacat cgaggaggtc gaggcggcgg gcggcatggc gcagggccatc 1320
173 gacgcgggca tcccgaagct gcgcgtcgag gagggcgcg cgcgacacca ggcgcgcatc 1380
174 gactcggggc gccagccggt catcggcgctc aacaagtacc ggggtggacac cgacgagcag 1440
175 atcgacgtcc tgaaggctga caactcctcg gtgcgcgccc agcagatcga gaagctgcgg 1500
176 cgctgcgcgc aggagcgtga cgacgccgcc tgccaggacg cgctgcgcgc cctgacggcg 1560
177 gcggccgagc gtggtcccgc ccagggactg gagggcaacc tgctcgcgct cgcggtcgac 1620
178 gcggccggcg cgaaggccac ggtcggtag atctccgac cactggagag cgtgtacggg 1680
179 cggcacgcgc gccagatccg tacgatctcc ggtgtgtacc gcaccgaagc aggccagtcg 1740
180 ccgagcgtgg agcgcacgcg tgccctggtg gacgcgttcg acgaggccga ggggcgcagg 1800
181 ccgcgcaccc tcgtcgcgaa gatgggtcag gacggccacg accgcggtca gaaggtgatc 1860
182 gcgagcgctt tcgcccacct gggcttcgac gtcgacgtcg gcccgtgtt ccagacgccg 1920
183 gcggaggtcg cgcccgaggc cgtcgaggcg gacgtgcaca tcgtcggcgt ctccctcgctt 1980
184 gccgcagggc acctaccctt cgtaccggca ctgcgcgagg agctggccgc ggagggccgc 2040
185 gacgacatca tgatcgctgt gggcgcgctc atcccgcgc aggacgtcga ggccctgcac 2100
186 gaggcgggcg ccacggcggt gttcccgcgc ggcacggtga tcccggacgc ggcgcacgac 2160
187 ctggtgaagc gtctggcggc cgacctcggc cacgaactgt ga 2202

```

189 <210> SEQ ID NO: 5

190 <211> LENGTH: 447

191 <212> TYPE: DNA

192 <213> ORGANISM: Propionibacterium freudenreichii ssp. shermanii

194 <400> SEQUENCE: 5

```

195 atgagtaatg aggatctttt catctgtatc gatcacgtgg catatgcgtg ccccgacgcc 60
196 gacgaggctt ccaagtacta ccaggagacc ttcggctggc atgagctcca ccgcgaggag 120
197 aacccggagc agggagtcgt cgagatcatg atggccccgc ctgcgaagct gaccgagcac 180
198 atgaccagcgg ttcaggatcat ggccccgctc aacgacgagt cgaccgttgc caagtggctt 240
199 gccaaagcaca atggctcgcg cgactgcac cacatggcat ggcgtgtcga tgacatcgac 300
200 gccgtcagcg ccaccctgcg cgagcgcggc gtgcagctgc tgtatgacga gcccaagctc 360
201 ggcaccggcg gcaaccgcat caacttcatg catcccaagt cgggcaaggg cgtgctcatc 420
202 gagctcacc agtaccgaa gaactga 447

```

204 <210> SEQ ID NO: 6

205 <211> LENGTH: 148

206 <212> TYPE: PRT

207 <213> ORGANISM: Propionibacterium freudenreichii ssp. shermanii

209 <400> SEQUENCE: 6

210 Met Ser Asn Glu Asp Leu Phe Ile Cys Ile Asp His Val Ala Tyr Ala

211 1 5 10 15

212 Cys Pro Asp Ala Asp Glu Ala Ser Lys Tyr Tyr Gln Glu Thr Phe Gly

RAW SEQUENCE LISTING

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:41

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw

213		20		25		30										
214	Trp	His	Glu	Leu	His	Arg	Glu	Glu	Asn	Pro	Glu	Gln	Gly	Val	Val	Glu
215			35				40					45				
216	Ile	Met	Met	Ala	Pro	Ala	Ala	Lys	Leu	Thr	Glu	His	Met	Thr	Gln	Val
217		50					55					60				
218	Gln	Val	Met	Ala	Pro	Leu	Asn	Asp	Glu	Ser	Thr	Val	Ala	Lys	Trp	Leu
219	65					70					75				80	
220	Ala	Lys	His	Asn	Gly	Arg	Ala	Gly	Leu	His	His	Met	Ala	Trp	Arg	Val
221				85						90					95	
222	Asp	Asp	Ile	Asp	Ala	Val	Ser	Ala	Thr	Leu	Arg	Glu	Arg	Gly	Val	Gln
223				100					105					110		
224	Leu	Leu	Tyr	Asp	Glu	Pro	Lys	Leu	Gly	Thr	Gly	Gly	Asn	Arg	Ile	Asn
225			115					120					125			
226	Phe	Met	His	Pro	Lys	Ser	Gly	Lys	Gly	Val	Leu	Ile	Glu	Leu	Thr	Gln
227		130					135						140			
228	Tyr	Pro	Lys	Asn												
229	145															

VERIFICATION SUMMARY

DATE: 08/29/2003

PATENT APPLICATION: US/09/699,136A

TIME: 14:19:42

Input Set : A:\010025-01 SEQ LISTING.DBF.TXT

Output Set: N:\CRF4\08292003\I699136A.raw